UCRL-18658

eyI

ABIOBENIC INFORMATION COUPLING BETWEEN NUCLEIC ACID AND PROTEIN, OR, HOW PROTEIN AND DNA WERE MARRIED

Melvin Calvin

December 1968

LAWRENCE RADIATION LABORATORY
UNIVERSITY of CALIFORNIA BERKELEY

UCRL-18658

eyI

Proceedings of the Royal Society
of Edinburgh

UCRL-18658
Preprint

UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

AEC Contract No. W-7405-eng-48

ABIOGENIC INFORMATION COUPLING BETWEEN NUCLEIC ACID AND PROTEIN, OR, HOW PROTEIN AND DNA WERE MARRIED

Melvin Calvin*

December 1968

UCRL-18658

ABIOGENIC INFORMATION COUPLING BETWEEN NUCLEIC ACID AND PROTEIN, OR, HOW PROTEIN AND DNA WERE MARRIED*

Melvin Calvin
Honorary Fellow, Royal Society of Edinburgh
and F.R.S.**
Physical Chemistry Laboratory, Oxford University,
Oxford, England

INTRODUCTION

The title to the talk this afternoon, "Abiogenic Information Coupling, or How Protein and DNA Were Married", is, indeed, an excerpt from the much broader subject which your president referred to, namely, chemical evolution of life on the surface of the earth, with some speculation about elsewhere as well. Actually, the "elsewhere" portion is more than a hopeful addition because it may provide a check on the speculations on the origin of life on the surface of the earth, which is very difficult to obtain in any other way.

I will not try to outline the whole sequence of events; most of you, I am sure, are familiar with parts of it.[1] I will try to pinpoint that particular step(s) in this chemical evolutionary sequence which now seems to me -- and I think to some of the Fellows here -- to be one

---

*Address to the Royal Society of Edinburgh, June 17, 1968

** Eastman Professor, Oxford University, Oxford, England, 1967-1968. Permanent address: Laboratory of Chemical Biodynamics, University of California, Berkeley. The preparation of this paper was sponsored, in part, by the United States Atomic Energy Commission.

of the crucial steps for whose essential character we have really not yet derived any basic concept and which is one of the blockades in the thinking, both of the biologists, on the one hand, and the mathematicians and philosophers, on the other, in trying to construct some foundation for believing that the origin of life on the surface of the earth is a consequence of the operation of the laws of physics and chemistry with no need to call upon other events than those in order to understand evolution in modern scientific terms.

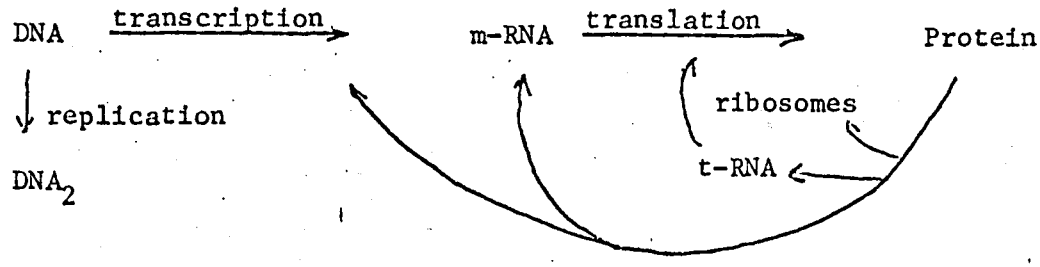
The principal agent for the construction of living material, both in the material form and in the way in which the design is transmitted from one generation to the next, resides in two essentially linear polymeric materials. These are two essentially one-dimensional molecules -- molecules which are similar to tapes, or strings, with a series of letters along them and which are informationally related to each other. Fig. 1 shows the two linear polymers of nature, protein and nucleic acid. What I mean by a linear array is that the substance(s) is a single chain which differs only in periodic points along that chain, and the information is stored in the array of those particular points. In the nucleic acid chain, the letters A,C,G and T are the shorthand letters for the four principal bases which constitute the composition of the nucleic acid--adenine, cytosine, guanine and thymine. The protein, on the other hand, which is really the major apparatus of a living organisms (controlling its function in terms of structure, material and energy manipulation) also shows a linear periodic difference, but the number of such different groups which are arranged along this periodic backbone is considerably larger, of the order of twenty or more. When this linear array

of the twenty different units some hundreds of units in length is finally put together, it takes on a rather complex three-dimensional structure dependent upon the particular linear array. This permits the enormous variety of chemical capability which these molecules must have in order to provide the necessary functions for life to proceed, i.e., manipulation of energy and material in a directed way.

The various components of both of these chains (protein and nucleic acid) have been constructed by abiogenic methods,[2] and we now have every basis for believing that we can envisage a series of chemical events, beginning with the naked earth, which would give rise to at least components of those linear arrays.

I would like to remind you very briefly of the way in which today's living organism makes the translation from the nucleic acid into the protein. Most of you have seen this representation at one time or another. The sequence of events is rather complicated, but one which is now in its major aspects fairly well described. Fig. 2 shows the presently operating coupling system in living organisms -- how the translation from the nucleic acid into the protein is achieved. We now know that it takes three of the letters of the DNA chain to designate one of the R groups on the protein chain, and the mechanism by which that is done is shown diagramatically. Some small segment of the desoxynucleic acid in the cell nucleus is transcribed onto a other kind of nucleic acid, which can move out of the nucleus into the cytoplasm and carry with it the necessary information from the DNA as a small template messenger-RNA. There, using this particular information for a particular protein, the assembly machinery goes to work and puts the particular

the assembly machinery goes to work and puts the particular amino acids, represented in Fig. 2 by lines of different length, together to create a protein. A simplified representation of this same process is shown below:



We achieve construction of a particular protein by hanging the amino acid onto a piece of transfer-RNA which has in it the correct triplet code for that particular amino acid. The matching up of the triplet code letters of the loaded t-RNA with the corresponding triplet on the messenger-RNA puts the amino acids in the right sequence. This is, as you can see, a rather complicated way of translating the information contained in one linear array into the construction of another linear array.

The question is one which was pointed out most explicitly just recently by Professor C. H. Waddington in Edinburgh: [3] How could this relationship, which is a very complex one, have arisen in the first place? This is the problem to which we are addressing ourselves this afternoon -- abiogenic information coupling between nucleic acids, on the one hand, and the protein, on the other. How did a relationship arise which is present in this developed form in all living cells today? You can see that there is a very special property in the t-RNA; it is the molecule that relates a particular code letter with a particular amino acid. I will now try to discuss some experimental evidence for the evolutionary appearance of that particular kind of relationship.

DEHYDRATION CONDENSATION: THE METHOD OF CHOICE

All of the components of the linear arrays (amino acids, bases, sugars, lipids) have long since been constructed by nonbiological methods,[2] that is, by radiation inputs (ultraviolet, electrical discharge, thermal, etc.). How can we chemically "grow" these units into linear arrays of nucleic acid and protein? Fig. 3 is an example of the kind of processes that had to be developed to generate the macromolecules. It shows the construction process for the various kinds of molecules. We have already indicated how we can make the small components (amino acids), and in order to hook them together we have to eliminate water between the amino group at the end of one and the carboxyl group at the end of the other, to form the polypeptide. This is a dehydration condensation which is the reaction required for the creation of the macromolecular species (polysaccharides, fats, and even nucleic acids; Fig. 4).[4, 5, 6]

Three different dehydration condensations are required to produce the linear array of the nucleic acid whereas only one is required to produce the linear array of the protein.

In every case, however, a dehydration condensation is required to create these macromolecular species, and it was necessary to determine how to induce this type of condensation by abiological methods. There have been a variety of ways developed to achieve this, and one of them is simply by heating up a mixture of amino acids and driving out the water, thus allowing the condensation to occur. This

is done under special circumstances, usually in a melt of one of the amino acids in excess -- glutamic acid is the usual one that is used. Table 1 shows the results of such a condensation starting with two moles of aspartic acid, one mole of glutamic acid and three moles of a mixture of all other amino acids in equal amounts.[7] When that mixture is heated at 100°C for 150 hours, there is created a polypeptide that is isolated from such a reaction mixture and which does not have the composition with which the mixture started.

TABLE 1

Thermal Condensation Reactions Showing Selectivity

Starting Mixture			Proteinoid Components		
			Product		
Aspartic acid	2	30%		52%	
Glutamic acid	1	15%		13%	
All others	3	55%			
			Lysine	2.8%	
			Histidine	2.5%	7.1%
			Arginune	1.8%	
			11 Others	27.8%	

You will note that the reaction started with aspartic acid (30%), 15% glutamic acid and 55% all other amino acids, and the product is 52% aspartic acid, 13% glutamic acid, 7% of the basic amino acids and only 27.8% of the remaining acids. This shows a rather widespread

difference between the ease with which these different amino acids get in. The purpose of Table 1 is to show that even when such violent dehydration conditions are provided (100°C for 150 hours) there is a certain degree of selectivity amongst the different amino acids. Some are destroyed and others are more selective and reactive, and a product results which is not simply the random collection of starting materials. There is a certain selectivity, or order, of amino acids that one gets in this collection and it is not a random order.

If we devise more sensitive, delicate methods of hooking the amino acids together in a mixture, we find that this selectivity can become even greater. Since this condition of heating may not have been a likely environment for the primitive earth, we undertook to find more likely conditions in water solutions. Here we made use of the multiple bonded carbon-nitrogen compounds which store a good bit of the electrical, or ultraviolet, energy that is thrown into the primitive atmosphere and creates them. One can use this multiple bond between carbon and nitrogen to extract the water between the various components of the mixture with which we are dealing. Fig. 5 shows how HCN could be used for this purpose (hooking together two amino acids to form the peptide while the water is taken out as formamide).[1a] To accomplish these reactions we now use cyanide derivatives which are found in the primitive discharges and which are related to one of the principal reagents which chemists use to hook together amino acids and other dehydration condensations, even in aqueous medium.[4, 8] These are the cyanamide derivatives shown in Fig. 6.[8] Fig. 6 demonstrates the use of cyanamide in dehydration condensation reactions of this type, through the carbodiimide struc-

ture produced by tautomerism. This is the fundamental structure which has been used to connect amino acids and other compounds in dehydration condensation reactions. As you can see, here is a way of coupling the amino acids in water in a relatively mild reaction. This has been done with a variety of amino acids, and Fig. 7 shows the kinetics on glycine polymerization by continuous addition of the condensing agent at a constant rate, showing the creation of some of these materials.[9]

SELECTIVITY BY GROWING POINT

I also want to bring to your attention the fact that there is a selectivity possible in this type of reaction as well. The way the selectivity was determined is as follows. Instead of mixing two amino acids, adding the condensing agent and allowing the amino acids to be hooked together and then examining the reaction mixture to determine the various products (four different possible combinations), one end of an amino acid is hooked to the wall, so to speak, and then tested with a series of others to see how fast they would hook onto the tail that was hanging loose. This gave a whole series of coupling rate constants of one amino acid (the one hung to the wall) with the others coupling to it in a specific way. We then begin to see a specificity in the coupling reaction. We suspected it should be there, and then it was actually found. Table 2 shows some of the selectivity that was found.[10] Using the rate of coupling of glycine onto glycine as unity (the norm) the rate of coupling of glycine to alanine, of alanine to alanine, etc., is expressed relative to it. You can see there is a widespread difference (a factor of 10) in the rates, the

TABLE 2

Comparison of Experimentally Determined Dipeptide Yields and Frequencies Calculated from Known Protein Sequences [10]

Dipeptide*	Values (Relative to Gly-Gly) Experimental	Calculated
Gly-Gly	1.0	1.0
Gly-Ala	0.8	0.7
Ala-Gly	0.8	0.6
Ala-Ala	0.7	0.6
Gly-Val	0.5	0.2
Val-Gly	0.5	0.3
Gly-Leu	0.5	0.3
Leu-Gly	0.5	0.2
Gly-Ile	0.3	0.1
Ile-Gly	0.3	0.1
Gly-Phe	0.1	0.1
Phe-Gly	0.1	0.1

* The dipeptides are listed in terms of increasing volume of the side chains of the constituent residues. Gly = glycine; Ala = alanine; Val = valine; Leu = leucine; Ile = isoleucine; Phe=phenylalanine. Example: Gly-Ala = glycylalanine.

phenylalanine to glycine being the slowest shown. In fact, phenylalanine-phenylalanine was so slow that the rates could not be included on the table. You can see, therefore, that there is this kind of specificity between the amino acids. We can then ask: Is there any sign that the same type of specificity has any influence on the structure of natural proteins?

We then look for the frequency of occurrence of those couplings (glycine-glycine, glycine-phenylalanine, etc.) in natural proteins. Using the reference work entitled the "Atlas of Protein Structure" published yearly,[11] one can count the couplings (doublet) in the naturally-occurring proteins and relate them all to the common one, for the glycine-glycine. You will see in the tabulation that the experimental doublets are actually in the right order (taken from the 1966-1967 "Atlas") with respect to the naturally-occurring couplings.

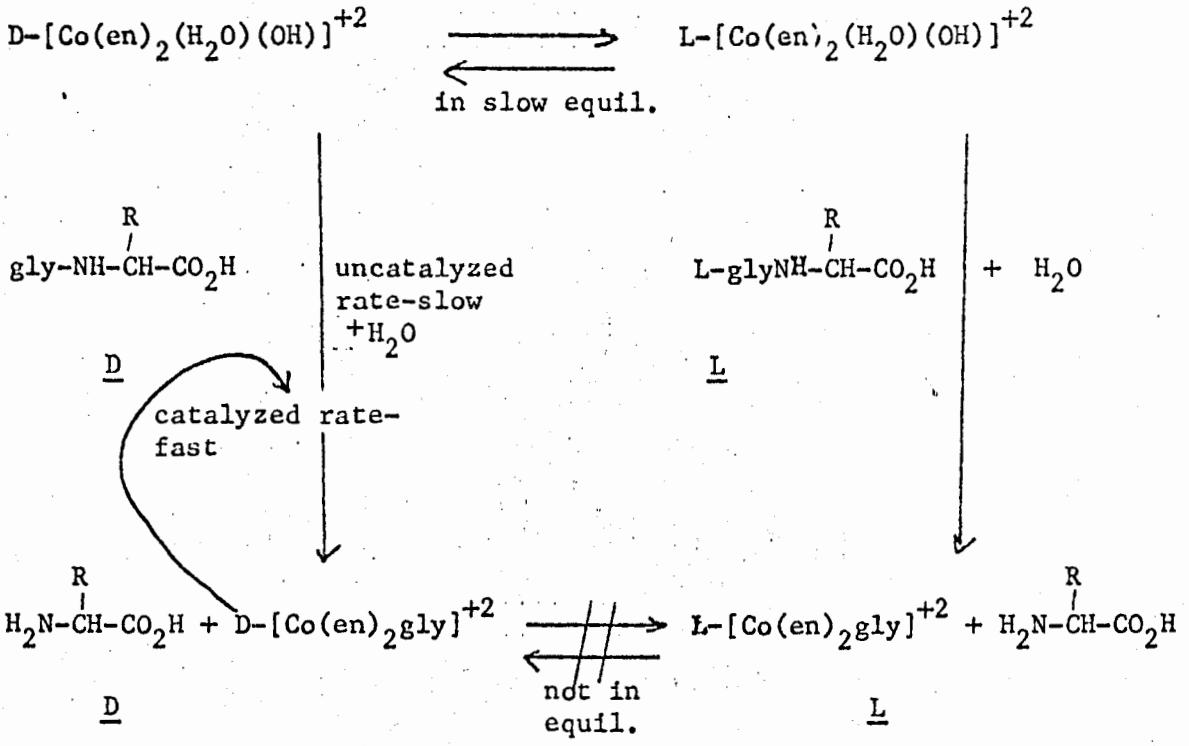
This lent support to the notion that sequences that we now find in protein may have some remnants left in them of what were the original sequences which were put in, not by the DNA mechanism, which we now know is the way they are put in, but by the nonbiological mechanism to start with which has been greatly modified during the course of evolution. There seems to be some evidence that a kind of selectivity was present in the primitive process, and is still visible even after more than 3 billion years of evolution has changed and modified the coupling mechanisms.

This gives, therefore, a kind of selectivity by the "growing point", probably influence by secondary folding, although the experi-

ments have not yet reached that level. One can imagine, now, a peptide chain growing by this coupling mechanism, and all the amino acids will not come onto that chain with equal ease and frequency. The "growing point" will be determined, first, by what is already present (not necessarily just the end, but some of the things that are behind that end, depending upon how far back that influence reaches, which is still not known) and also by the distribution of units available to put onto the chain. The two factors will, together, determine what comes on next in such a "growing" material. However, this still does not give a method for replicating any particular sequence.

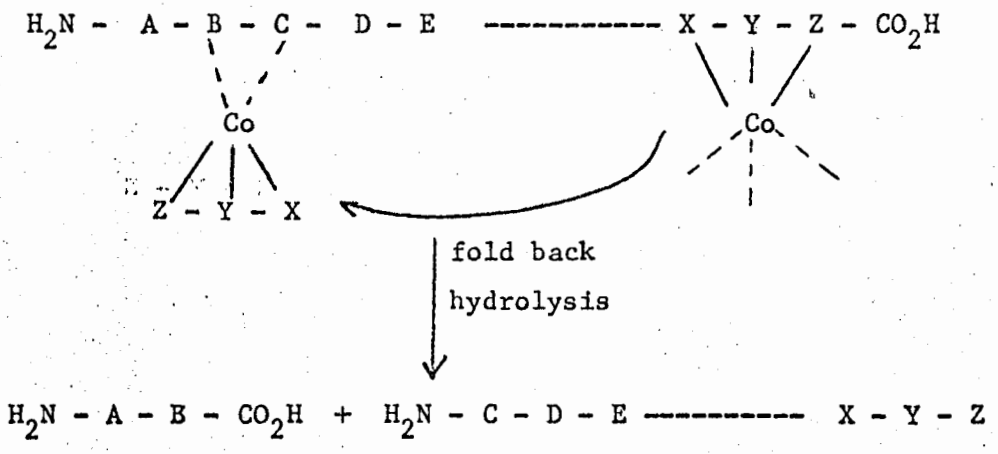
REPLICATION METHODS

There should be some way of replicating a particular sequnce, and I have tried to invent one, using some recently described catalytic mechanisms for growing a polypeptide, for chopping off the end, and then to start it growing again. How can this be done in a specific manner? It can be done as follows: The cobalt complex, bis-ethylene-diamine hydroxy aquo cobaltous ion, which is octahedral in shape and optically active, is in equilibrium in the D and L forms. Such a cobalt complex can react with a peptide which will be hydrolyzed, the acidic end group of the peptide being dropped off, and an optically-active cobalt complex resulting therefrom. This complex can conceivably then react with another peptide, and make a very fast catalyzed reaction compared to the uncatalyzed slow reaction with which the sequence began. The initial equilibrium is slow compared to the catalyzed reaction, and all the material will go in the initially catalyzed direction.[12]



Thus, whichever series starts first will determine the optical character of both the liberated acid and the remaining peptide.[13] Once you get the D-cobalt complex, only the D complex will continue; this is an optical selectivity which is even a higher degree of selectivity than amino acid selectivity itself. The nature of the R-group will surely effect the rate of this reaction. This is simply a method for breaking off a peptide link somewhere near the end and one can use that method to make a crude kind of coding replication which does not involve the present-day nucleic acid sequencing. Below is shown a model system, in which the letters A,B,C,D,X,Y,Z represent particular peptide links. Let us suppose that one particular end has some specific amino acids, onto which the cobalt can be attached, and that the bound

Co can reach back and induce hydrolysis between a particular pair of amino acids, just as in the reaction demonstrated above. This would give a small dipeptide of a specific character which can regrow in the original way.[14] So, this is a crude kind of protein replication,



which, obviously is not very "hi fi". What you would get the next time you put the material through the "growth" process will not be identical with what you started with; it will merely resemble it to a degree, depending upon the conditions of the environment and the raw materials which are available. There is a certain probability of specific replication of peptide. Keep in mind, now, that the peptide has all of the catalytic capacity because of its ability to have the enormous variety of structure, even though it does not have a very good replicating capacity. We have generated here a scheme for making a whole set of more or less particular polypeptide structures and regenerating them, in a

crude and not very "hi fi" manner. Nevertheless, it is a protein replication system which does not involve the nucleic acids.

What about the nucleic acid system? Can we generate the sequences there as well? We can generate polynucleotides even more easily by exactly the same types of reactions that were used for the polypeptides -- the same dehydration condensation reaction can be used. Three different kinds of dehydration condensations are required to generate a polynucleotide, and the same molecules (cyanamide, dicyanamide, polyphosphate) as mentioned before can be expected to do it. Here, also, there should be some selectivity between the four different bases, but the actual experiments have not yet been done. Therefore, I cannot say whether there is much of a selectivity in the purely chemical non-templated sequencing of the different nucleic acid components, although I suspect there will be a slight selectivity apparent in this as well as in the peptides. If you start with a random mixture of A,C,T and G, and even if you did not have base-pairing but simpling the coupling mechanism, the resulting polymer would probably not be completely random but some sequences would be more favored than others and you would get some kind of selectivity in much the same way as with the polypeptides.

The replication experiments on the polynucleotides have been done repeatedly and in several different ways. In contrast to the polypeptide this nucleic acid replication is a very "hi fi" reproducing system even without any enzymes present -- abiologically, in other words. It is also relatively easy to accomplish, and Fig. 8 shows the results of Schramm who polymerized uridine monophosphate in the presence of polymeric adenylic acid. The presence of polyadenylic acid enhances the

rate of uridine polymerization because of the uridine base-pairing with the adenine.[15] This shows that in the presence of poly-A the uridine polymerizes very well and in its absence the uridine polymerizes very slowly. This experiment, in which a polyphosphate anhydride is used as a condensing agent is analogous to the experiment of Fox's[7] which I described in which polyphosphoric acid and 100°C temperature were used to hook the amino acids together. The experiment of Schramm is the same kind of violent (vigorous) condensation, and the nature of the products is uncertain, to a degree. A much more elegant experiment by Naylor used carbodiimide as a condensing agent.[16] Fig. 9 shows the results of that experiment in which he demonstrated that, for example, hooking together a pair of hexamers of thymine in the presence of polyadenylic acid produced a reaction in which the poly-A catalyzed the reaction for the poly-T, and vice versa. The two things together form a reflexive catalytic system, one catalyzing the formation of the other. Thus, a replicating system for the nucleic acids can be created. This particular reaction happens to be very selective and "hi fi" for the nucleic acid and its fidelity has been demonstrated in the test tube without the use of biological agents.

We thus have two systems, separately devised. The first is the linear polypeptide system which has in it the multicatalytic capacity that we need. By arranging the polypeptides in a variety of orders and shapes it is quite clear that we can produce almost any catalytic function that a living organism might require. The fidelity of the replication of the linear polypeptide system is, however, very poor. It is a statistical replication, depnnding upon the selectivity of

the growing point and perhaps of some amino acids further back from the growing point and the availability of raw materials. The second system, that of the polynucleotides, is a very efficient and reliable replicating system, but it has no other function. It does have the high fidelity replicability that is not present in the polypeptides.

It is quite clear that the selective advantage of coupling these two qualities -- the catalytic quality of the protein and the replicating fidelity of the nucleic acid -- is what has to be achieved in order to give rise to the stability and improvement (evolution) of energy and material processing systems against the thermal disruptive forces of entropy; a stability with which we characterize a living thing.

MODEL EXPERIMENTS OF COUPLING REACTIONS

There are various ways in thich this kind of information storage and translation coupling has been sought. Not any of them yet have been very satisfactory (at least not to my mind), and one can imagine a variety of physical-chemical systems which would have some of these properties, particularly phase boundaries, surfaces, clay particles, etc. Unfortunately, there have been no experimental tests of most of these theoretical suggestions.[17] What I am going to do here is to say that we have ourselves thought of one way of coupling these two different types of reactions -- in other words, we have taken the next step and have done an experiment to test the viability of the notion as a step in the evolution of the coupled systems.

If you recall, I used the same kind of an idea in generating both of these linear sequences independently, namely, a type of "growing point" control. In other words, the shape and form of the growing

end of the nucleic acid (or polypeptide) and probably the conformation of the chain some way back had some effect on what was hooked to the end of it. The shape and form of the polypeptide would have some effect on what the next amino acid to be added would be. In the case of the polypeptides we were able to demonstrate approximately a ten-fold difference in selectivity under equivalent raw material conditions. In the case of polynucleotides, this information is not yet available.

The obvious thing came to us when we considered the way in which the present-day living organisms accomplish this coupling, which is by hanging a particular amino acid on a relatively short piece of nucleic acid. Somewhere in that nucleic acid there is a specific array of bases which is characterized as belonging to that particular amino acid. When we look at all of the carrier, or transfer-RNA's, which carry the amino acid over from the free state into the assembly machinery, we find that they end in the same three bases, -C-C-A. It occurred to us that perhaps this was an important clue. Perhaps the fact that all of the t-RNA's end in -C-C-A is a residue of some prebiological event, some residue of the chemical selectivity of -C-C-A for amino acids which any other base combination does not have. Thus, simplifying it still further, this suggests that if we could make a model of the t-RNA and try to couple different amino acids to it, we would have a selectivity in the amino acids for a particular base and we would also find some selectivity of the bases for the amino acids in the t-RNA analogy. That is exactly the study we have begun.

The simplest experiment was to make a model substance to represent the t-RNA and put adenine on the end of it. We therefore used the same polymer which was used to hang the amino acid on, i.e., polystyrene with a functional group on it.[18] Figs. 10 and 11 show the basic concept of such experiments. So far, we have only done the experiment using one base at a time, and we can now determine the efficiency, or the ease or the probability, of coupling amino acids to such a terminal nucleic acid end, and see if there is a difference in different bases. Figs. 12, 13 and 14 give a few details of the chemistry of this experiment. Fig. 12 shows how we hang the adenylic acid (A) on the resin by simply taking adenylic acid and the chlormethylated resin in pyridine for a few ours. Fig. 13 shows how we hang the amino acid onto the A, and there are several different ways in which this can be done. We have used the amino acid anhydride in pyridine and allowed it to react with the model t-RNA to get the product, for example. The details of taking the coupled product off the resin, using 80% acetic acid for 2.5 hours at room temperature, is shown in Fig 14, while Fig. 15 shows the other way of coupling, using the carbodiimide, which I mentioned earlier. We here are using the carbobenzoxyamino-protected phenylalanine (Z-Phe) to get the desired product.

Remember that I mentioned that we were going to try different bases on the resin as models for the transfer-RNA. So far we have only done two, with one base each time. We have tried to couple phenylalanine and glycine to these bases, with the results shown in Table 3.

TABLE 3

Percent of Bound Nucleotide Reacted

Base/Amino Acid	Adenine	Cytosine
Phenylalanine	6.7	2.9
Glycine	10.0	6.5

Quite clearly, adenine is a better acceptor than cytosine which is the first result of this experiment and which might have been expected. Furthermore, glycine goes onto the resin-bound nucleotide more readily than phenylalanine, which is also not unexpected. This is a degree of selectivity, which is what we are seeking, in the coupling of the amino acid to the base.[19] It is the beginning of the translation process, really. Remember, there are no enzymes present in this type of experiment (no biological material, i.e., no bacterium, no virus particle, no extract of a yeast, etc.) to "tell" the phenylalanine it should go on the adenine instead of the cytosine. This is a purely "chemical" translation. There is a selectivity in operating in this kind of a chemical translation process. Quite obviously, when we have a doublet base, C-A, we may expect a different degree of selectivity than we do with A alone. The selectivities will change, and as we make a doublet, triplet, and further on back, it is conceivable that we can elucidate the controlling structure which determines what hangs on the end and where it lies with respect to the end. As this work develops, I believe there will be certain parts of the molecule that reach back to the growing end which will help determine the sequence of events there. We will then have a high degree of selectivity which will actually be determined not merely by the growing end but by something quite far removed from that end.

CONCLUSION

There is now experimental evidence for selectivity between the amino acid and the nucleic acid base which is the beginning of the chemical translation process from one linear system to the other. The linear system of the nucleic acid is, of course, an excellent place to store the information, wheras the linear system of the polypeptide, on the other hand, is the versatile system which can perform many different types of reactions but is unable to store information reliably. The experiments I have described here may represent the beginning of the method of coupling of those two essential qualities which are required for the generation and evolution of a living organism.

REFERENCES

1. For reviews of the theory of chemical evolution see:

   a. Melvin Calvin. "Chemical Evolution", University of Oregon Press, Eugene (1961).

   b. Melvin Calvin. Chemical Evolution (The Bakerian Lecture). Proc. Roy. Soc. 288A, 441 (1965).

   c. S. W. Fox (ed.). "The Origins of Prebiological Systems and of Their Molecular Matrices", Academic Press, New York (1965).

   d. A. I. Oparin (ed.). "The Origin of Life on Earth (3rd English edition, trans. by Ann Synge), Oliver and Boyd, London (1963).

   e. J. D. Bernal. "The Origin of Life", Weindenson and Nicolson, London (1967).

   f. G. Steinman and D. H. Kenyon. "Biochemical Predestination", McGraw-Hill Publishing Co., New York (in press).

   g. Melvin Calvin. "Chemical Evolution", Oxford University Press, Oxford, England (1969).

2. For reviews of abiogenic synthesis of proteins and nucleic acids see:

   a. R. M. Lemmon. Abiogenic Synthesis of Biologically Relevant Organic Compounds. C hemical Reviews (in press).

   b. C. A. Ponnamperuma and N. W. Gabel. Current Status of Chemical Studies of the Origin of Life. Space Life Sciences, 1, 64 (1968).

3. C. H. Waddington. Towards a Theoretical Biology. Nature, 218, 525 (1968).

4. H. G. Khorana. "Some Recent Developments in the Chemistry of Phosphate Esters of Biological Interest", John Wiley and Sons, New York (1961).

5. Gary D. Steinman. Protobiochemistry: Dehydration Condensation in Aqueous Solution. PhD thesis, University of California, Berkeley, December 1965; University of California Lawrence Radiation Laboratory Report UCRL 16566.

6. G. Steinman, D. H. Kenyon and M. Calvin. Dehydration Condensation in Aqueous Solutions. Nature, 206, 707 (1965).

7. S. W. Fox. How Did Life Begin? Science and Technology, Feb. 1968, p. 61.

8. a. G. Steinman, R. M. Lemmon and Melvin Calvin. Cyanamide: A Possible Key Compound in Chemical Evolution. Proc. Nat. Acad. Sci. 52, 27 (1964).

b. A. Schimpl, R. M. Lemmon and M. Calvin. Formation of Cyanamide under "Primitive Earth" Conditions. Science, 147, 148 (1965).

c. G. Steinman, R. M. Lemmon and M. Calvin. Dicyandiamide: Possible Role in Peptide Synthesis during Chemical Evolution. Science, 147, 574 (1965).

9. D. H. Kenyon, G. Steinman and Melvin Calvin. The Mechanism and Protobiochemical Relevance of Dicyanamide-Mediated Peptide Synthesis. Biochim. Biophys. Acta, 124, 339 (1966).

10. G. Steinman and M. N. Cole. Synthesis of Biologically Pertinent Peptides under Possible Primordial Conditions. Proc. Nat. Acad. Sci. 58, 735 (1967).

11. R. V. Eck and M. O. Dayhoff. "Atlas of Protein Structure and Function, 1966." National Biomedical Research Foundation, Silver Spring, Maryland (1967).

12. D. E. Allen and R. D. Gillard. Stereoselective Effects in Peptide Complexes. Chem. Commun. 1967, 1091.

13. An actual experimental observation of spontaneous generation of optically active material which involves these principles of equilibrium and relative autocatalytic rates was made some fifteen years ago. E. Havinga. Spontaneous Formation of Optically Active Substances. Biochim. Biophys. Acta, 13, 171 (1954).

14. J. P. Collman and E. Kumura. The Formation of Peptide Bonds in Coordination Sphere of Cobalt (III). J. Am. Chem. Soc. 89, 6096 (1967).

15. G. Schramm, W. Grotsch and W. Pollman. Nonenzymatic Synthesis of Polysaccharides, Nucleosides, Nucleic Acids and the Origin of Self-Reproducing Systems. Angew. Chem. Int. Ed. 1, 1 (1962).

16. R. Naylor and P. H. Gilham. Study of INteractions and Reactions of Oligonucleotides in Aqueous Solution. Biochemistry, 8, 2722 (1966).

17. a. J. D. Bernal. The Problem of Stages in Biopoesis. In "The Origin of Life on the Earth", A. I. Oparin (ed.), Pergamon Press, London (1959), 38-53.

b. S. Akabori. The Origin of the Fore-Protein, ibid., 189-196.

c. A.G. Cairns-Smith. The Origin of Life and the Nature of the Primitive Gene. J. Theoret. Biol. 10, 53 (1966).

18. M. A. Harpold and M. Calvin. AMP on an Insoluble Solid Support. Nature, 219, 486 (1968).

19. M. A. Harpold and M. Calvin. Amino-Acid Nucleotide Interactions on an Insoluble Solid Support. I. A Simple Model of the Amino Acid Acceptor Terminus of a t-RNA.

FIGURE CAPTIONS

Figure 1 Structure of protein and nucleic acid (MU 22814)

Figure 2 A sequence of events in the molecular communication problem (MuB 938A

Figure 3 condensation
Dehydration/reactions of amino acids (MuB 5261)

Figure 4 Dehydration condensation reactions of nucleic acids (MuB 5262)

Figure 5 A mechanism of peptide formation with HCN as dehydrating agent (MU 27656)

Figure 6 A mechanism for peptide synthesis with carbodiimide and dicyanamide as dehydrating agents (BC 1053)

Figure 7 Rates of glycine utilization and polypeptide accumulation (MuB 8960) in solution containing glycine and HCl.

Figure 8 Polymerization of uridine monophosphate in presence (A) and absence (B) of polyadenylic acid. (Schramm, 1962; no number)

Figure 9 The coupling of two hexanucleotides to form a dodecanucleotide MuB 10896

Figure 10 Linear polymeric structure of nucleic acids and proteins (XBL 684 4138)

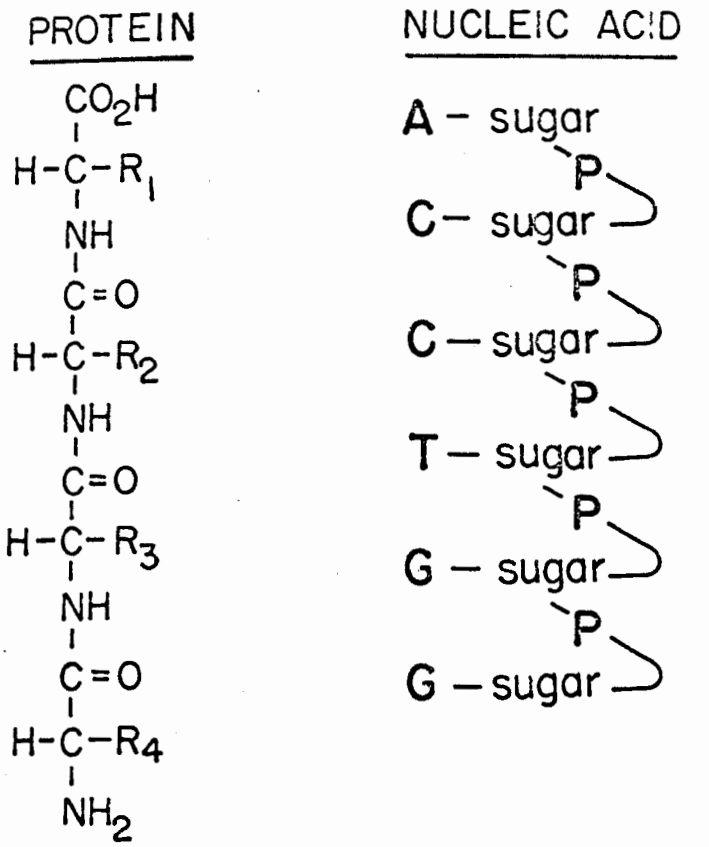
Figure 11 Base pairing for DNA replication and RNA template formation (XBL 684 4141)

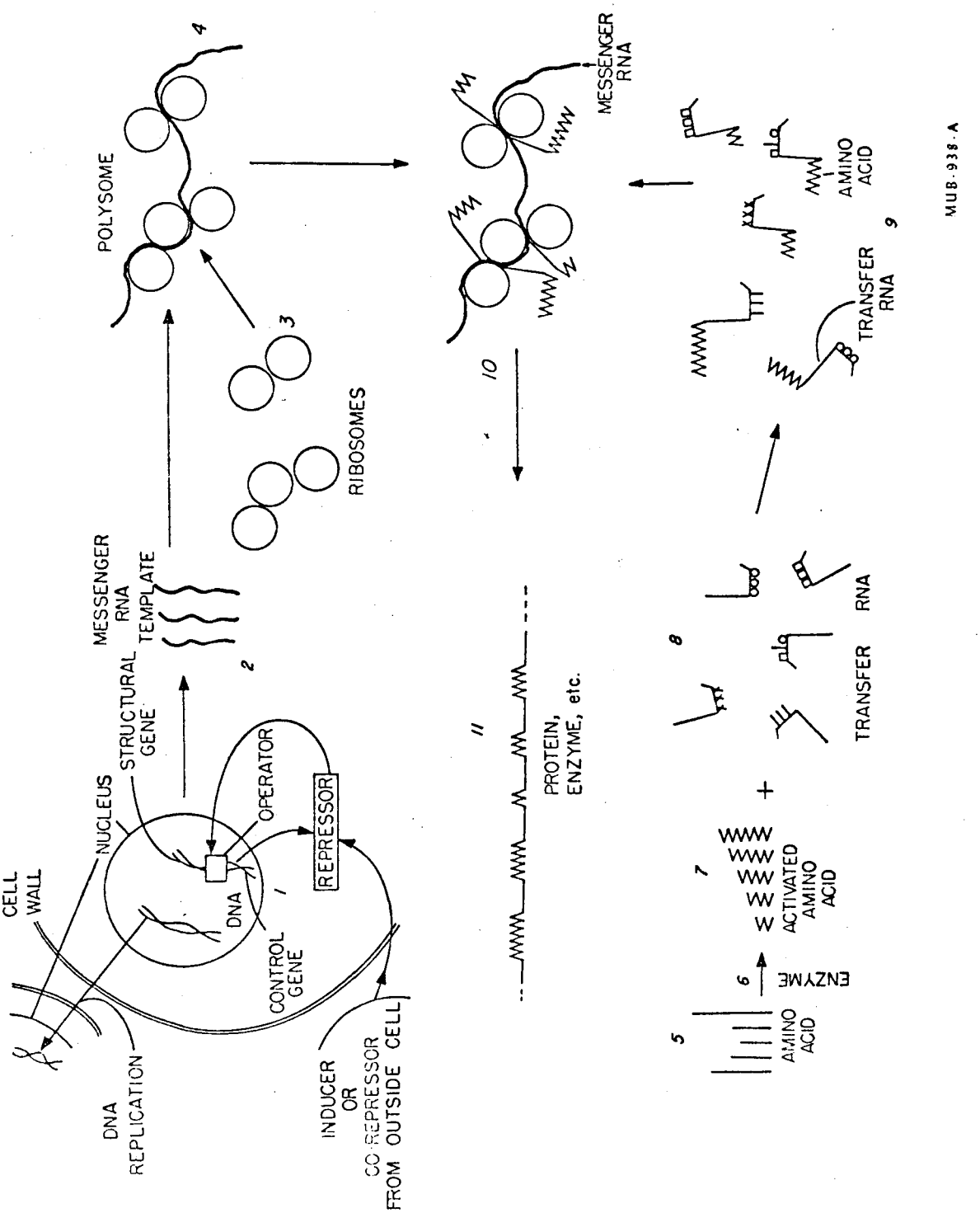
Figure 12 Coupling of AMP with the polymer and release of free AMP by hydrolysis (XBL 681 4009)

Figure 13 Coupling of polymer-AMP complex with the anhydride form of a N-protected amino acid (XBL 684 4137)

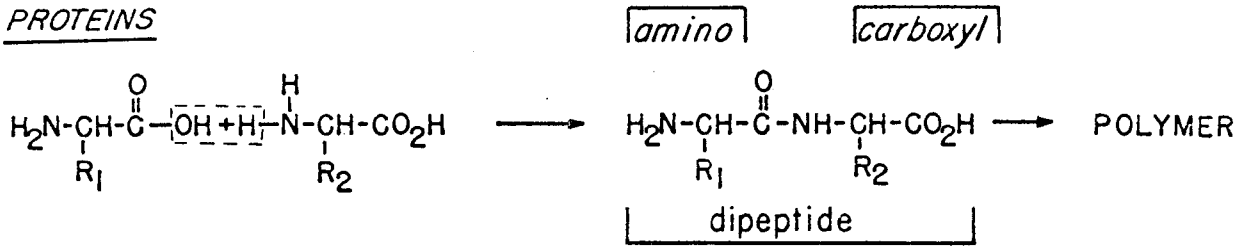
Figure 14 Another method of coupling an N-protected amino acid onto the polymer-AMP complex (XBL 681 4010)

Figure 15 Hydrolysis of the complex to release N-protected aminoacyl t-RNA (XBL 681 4004)

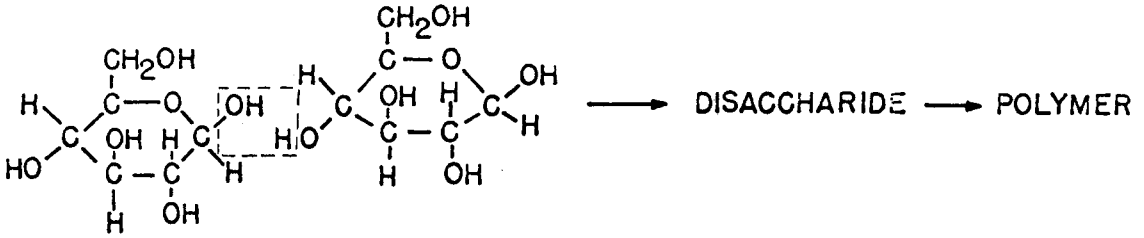

PROTEIN
NUCLEIC ACID
A – sugar
P
C – sugar
P
C – sugar
P
T – sugar
P
G – sugar
P
G – sugar
MU – 22814


CELL WALL
NUCLEUS
STRUCTURAL GENE
OPERATOR
REPRESSOR
DNA
CONTROL GENE
DNA REPLICATION
INDUCER OR CO-REPRESSOR FROM OUTSIDE CELL
MESSENGER RNA TEMPLATE
POLYSOME
RIBOSOMES
PROTEIN, ENZYME, etc.
MESSENGER RNA
AMINO ACID
TRANSFER RNA
ACTIVATED AMINO ACID
ENZYME
TRANSFER RNA
AMINO ACID
MUB-938-A

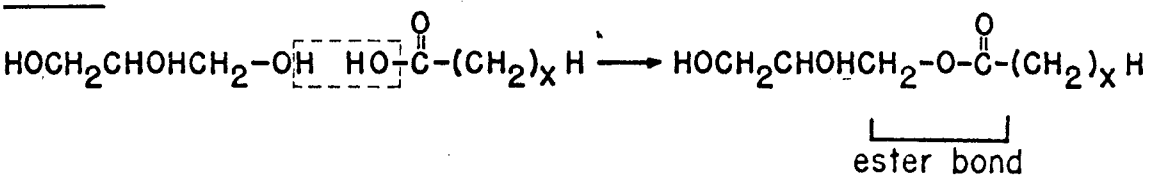
PROTEINS



POLYSACCHARIDES

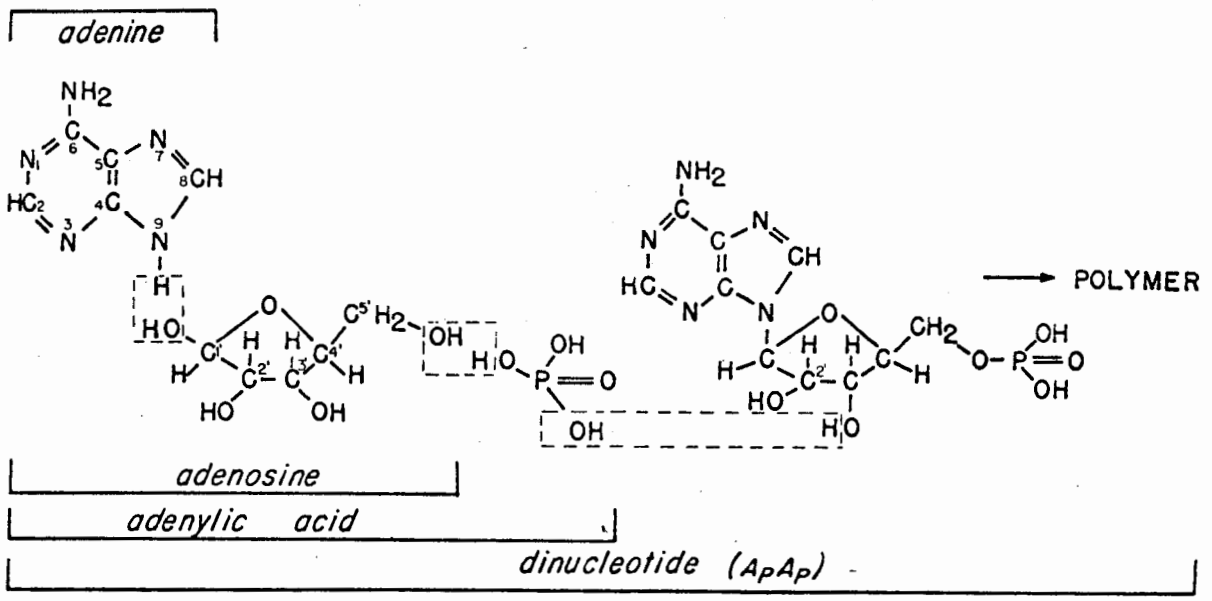


LIPIDS

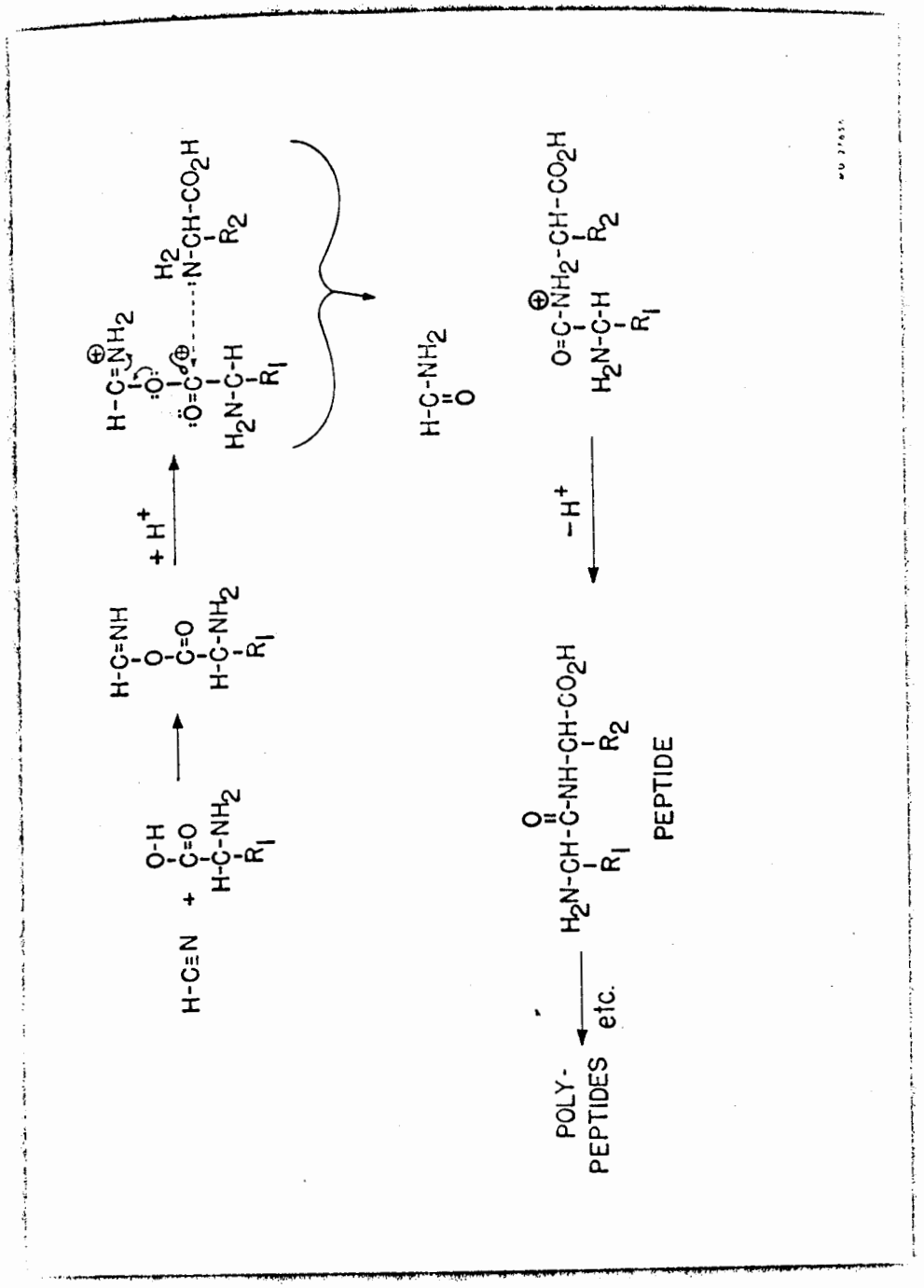


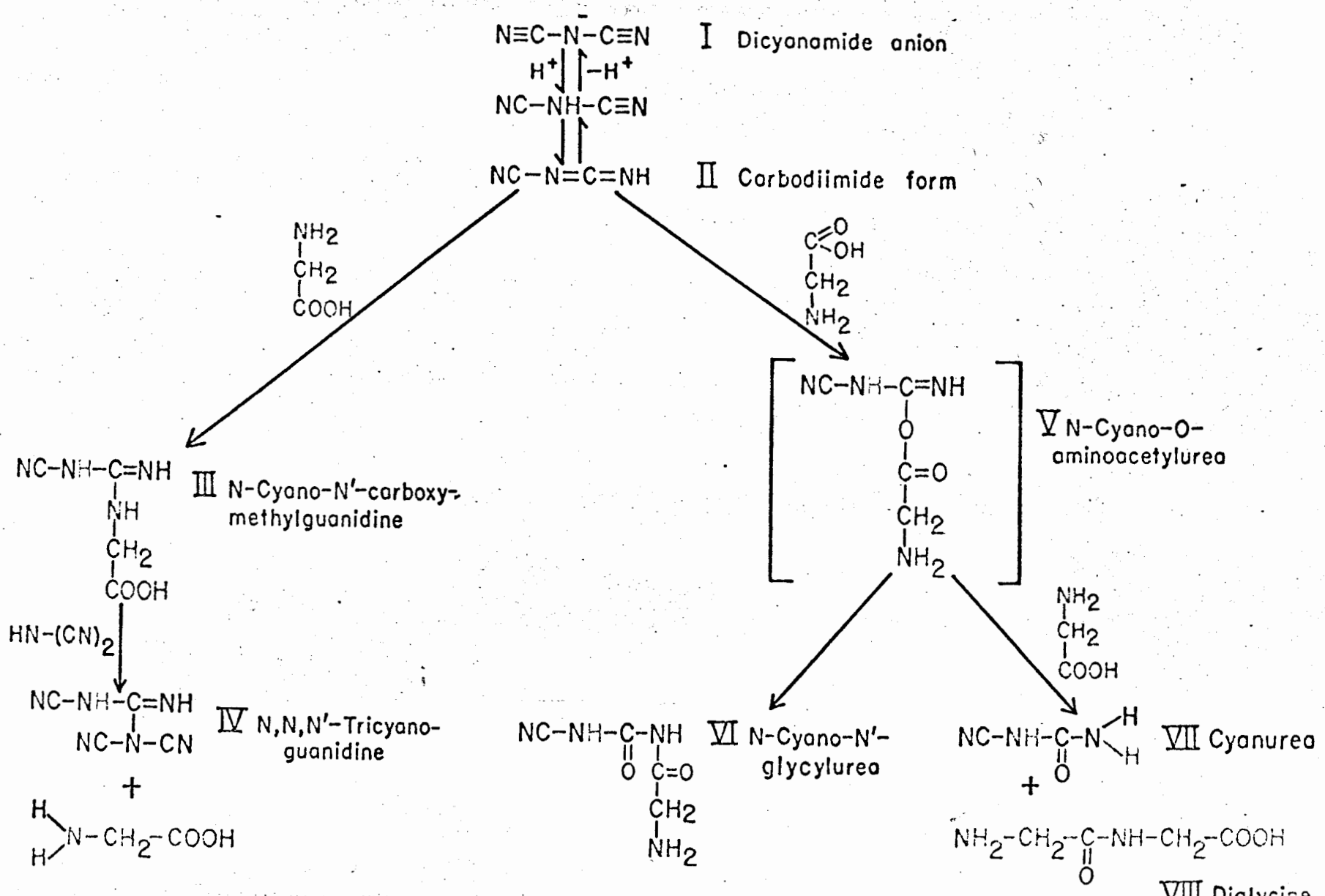
MUB-5261

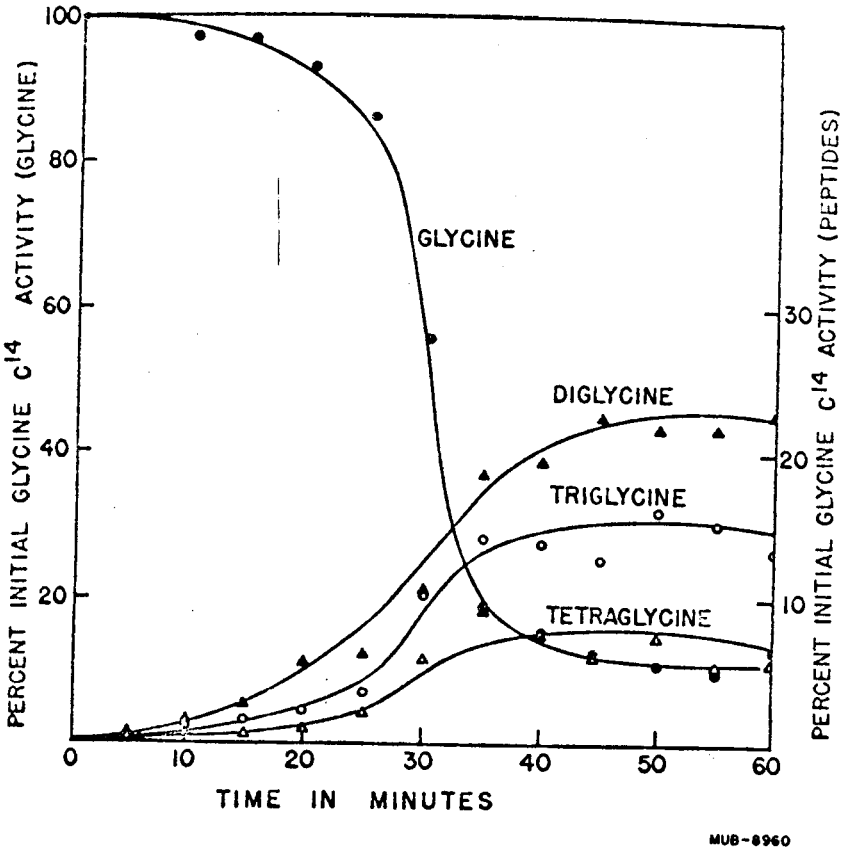
NUCLEIC ACIDS (3 STAGES) RNA SHOWN - DNA LACKS OH ON 2' POSITION

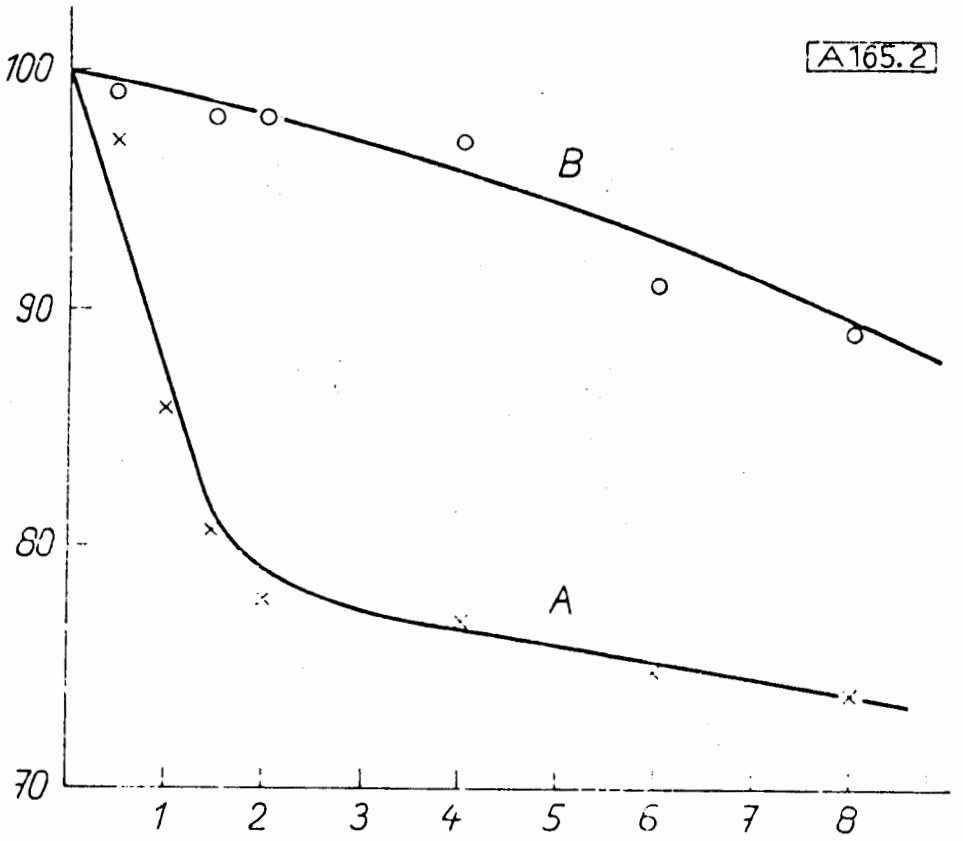


MUB-5262


H-C≡N
+
O-H
C=O
H-C-NH2
R1
H-C=NH
O
C=O
H-C-NH2
R1
+ H+
H-C=NH2
O:
:O=C
H-C-H
H2N
R1
H2
N-CH-CO2H
R2
H-C-NH2
O
O=C-NH2
H2N-C-H
R1
CH-CO2H
R2
-H+
H2N-CH-C-NH-CH-CO2H
R1
R2
O
PEPTIDE
POLY-
PEPTIDES
etc.


I Dicyanamide anion
II Carbodiimide form
III N-Cyano-N'-carboxy-methylguanidine
IV N,N,N'-Tricyano-guanidine
V N-Cyano-O-aminoacetylurea
VI N-Cyano-N'-glycylurea
VII Cyanurea
VIII Diglycine


100
80
60
40
20
0
10
20
30
40
50
60
30
20
10
GLYCINE
DIGLYCINE
TRIGLYCINE
TETRAGLYCINE
PERCENT INITIAL GLYCINE C14 ACTIVITY (GLYCINE)
PERCENT INITIAL GLYCINE C14 ACTIVITY (PEPTIDES)
TIME IN MINUTES
MUB-8960

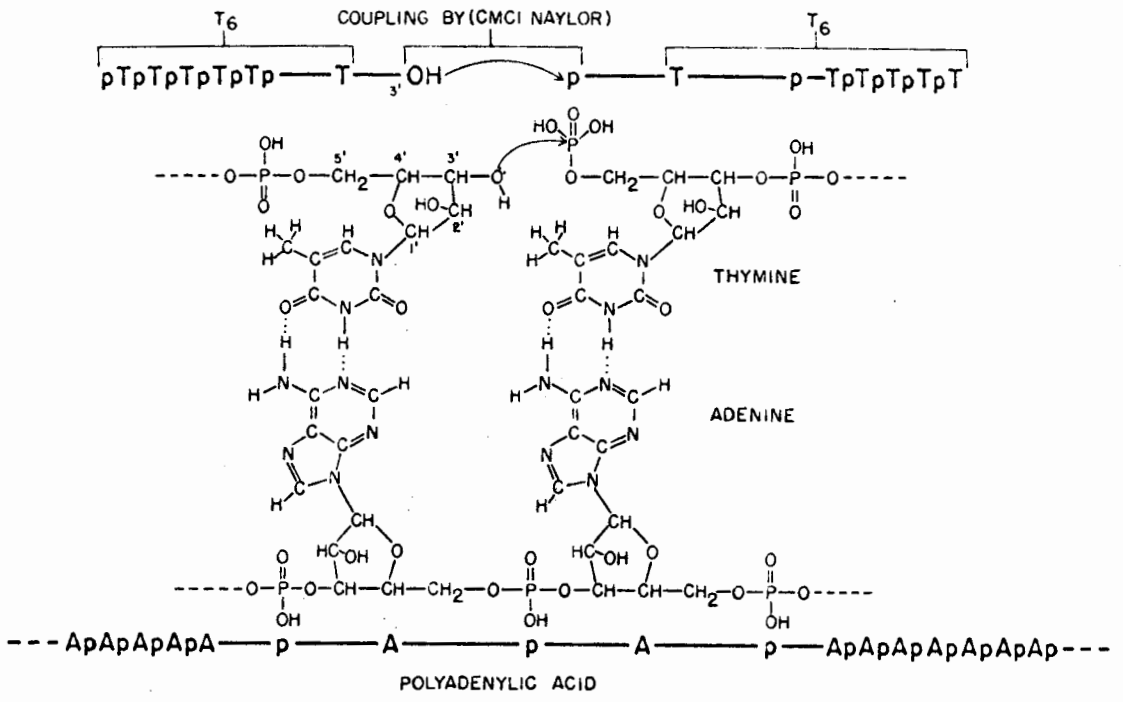


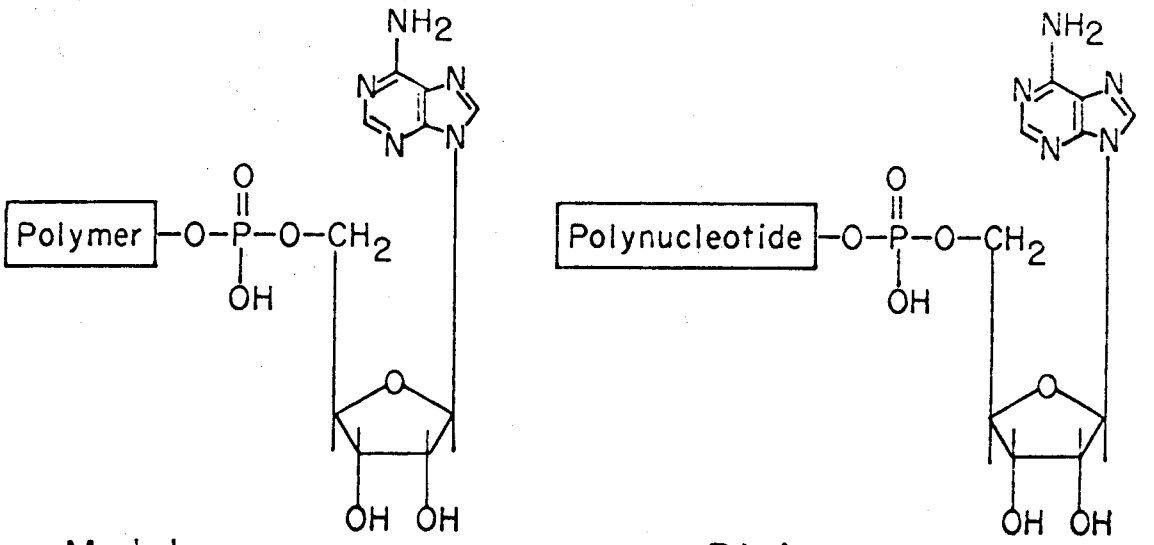
Polymerization of Uridine Monophosphate in the Presence (A) and in the Absence (B), of Polyadenylic acid. (The decrease of free uridine monophosphate was measured chromatographically).

Abscissa: Time [hours]

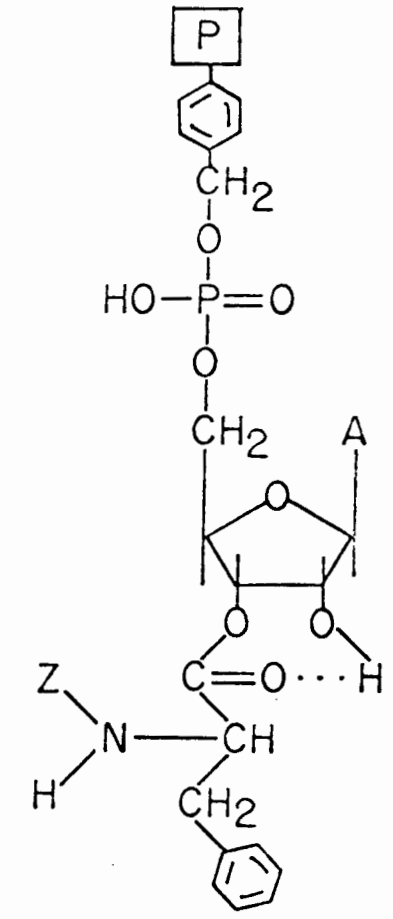
Ordinate: % Free Uridine Monophosphate (referred to the amount of starting material)

(Schramm, 1962)

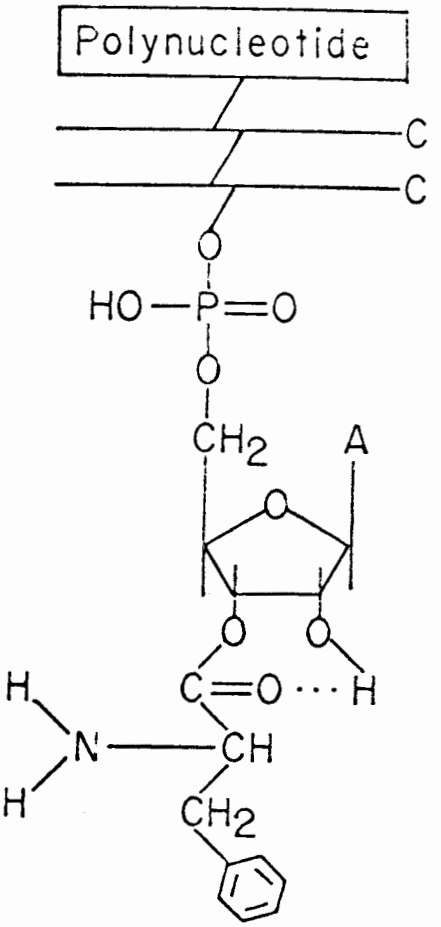

T6
COUPLING BY (CMCI NAYLOR)
T6
pTpTpTpTpTp
T
3'OH
p
T
p-TpTpTpTpT
THYMINE
ADENINE
---ApApApApA
p
A
p
A
p
ApApApApApApAp---
POLYADENYLIC ACID
MUB-10896


NH2
Polymer
O
P
O
CH2
OH
OH
OH
Model
Template
NH2
Polynucleotide
O
P
O
CH2
OH
OH
OH
tRNA
Terminus

XBL 684-4138

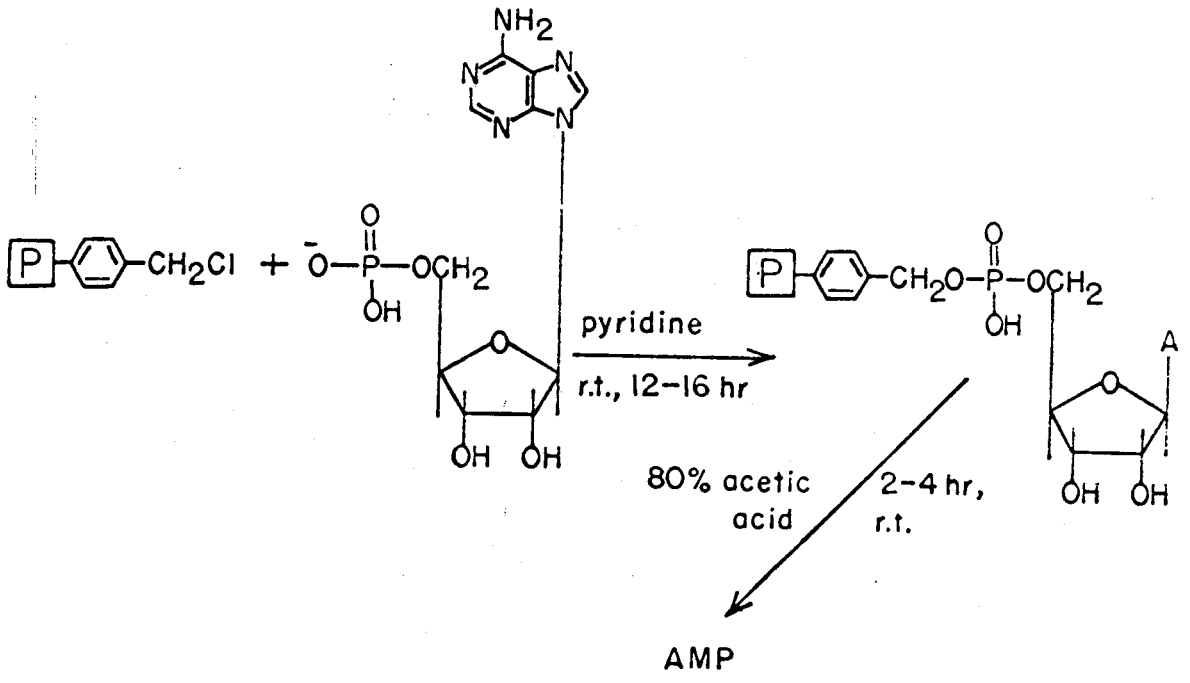


Model Phe tRNA

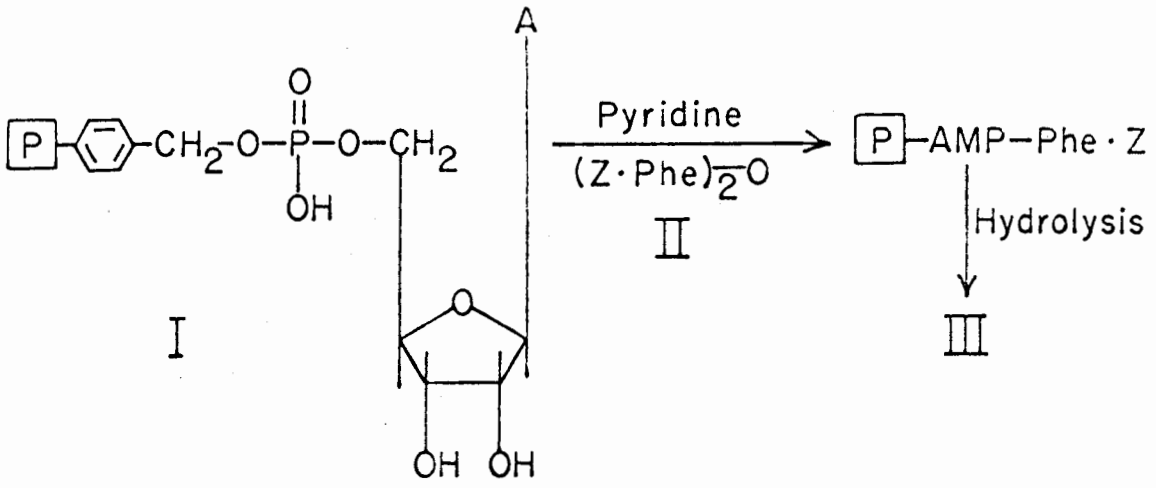


Phe tRNA

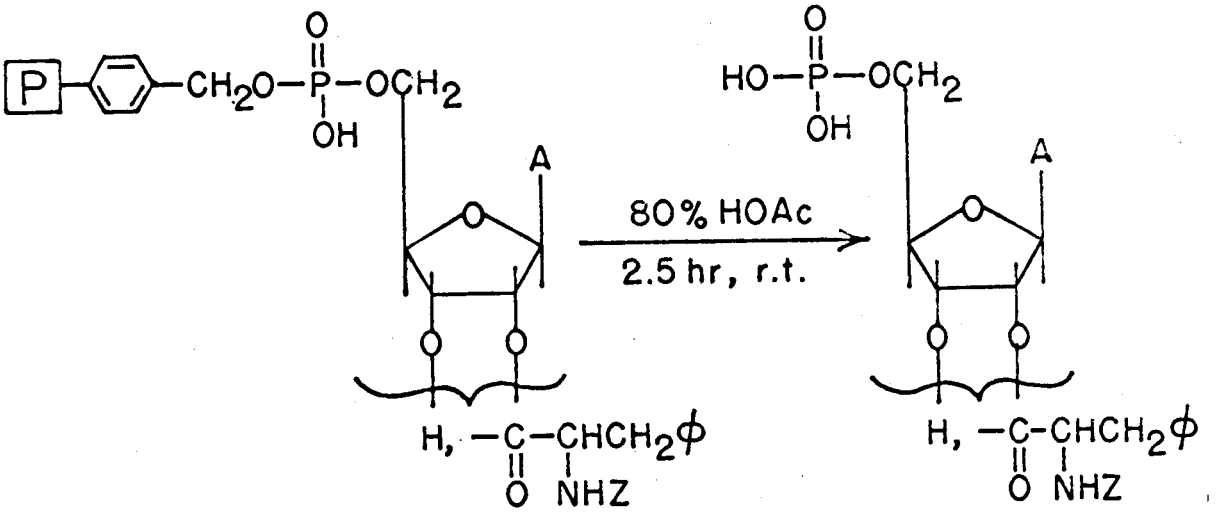
XBL 684-4141


NH2
P
CH2Cl
+
O
P
OCH2
OH
OH
OH
pyridine
r.t., 12–16 hr
P
CH2O
P
OCH2
OH
A
OH
OH
80% acetic acid
2–4 hr, r.t.
AMP

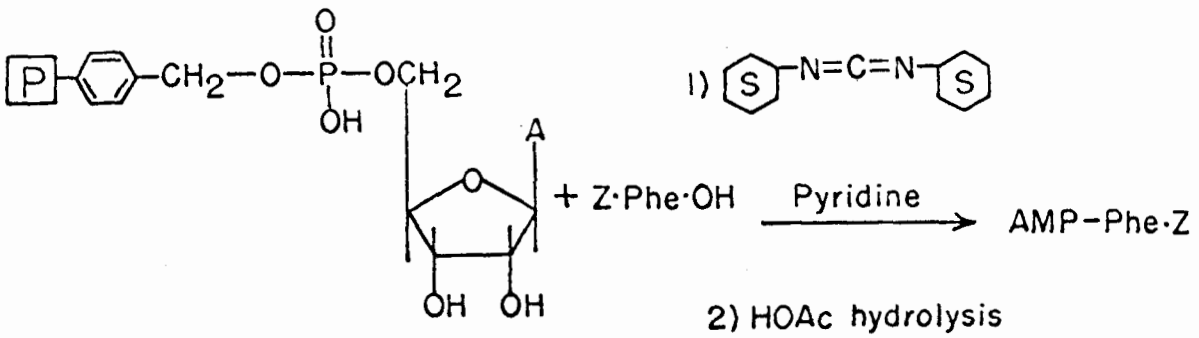
XBL 681-4009


A
P
CH2-O-P-O-CH2
O
OH
Pyridine
(Z·Phe)2O
P-AMP-Phe·Z
Hydrolysis
I
II
III
OH OH

XBL 684-4137



XBL 681-4010


P
CH2—O—P—OCH2
O
OH
A
OH OH
+ Z·Phe·OH
1) S N=C=N S
Pyridine
AMP–Phe·Z
2) HOAc hydrolysis

XBL 681-4004

LEGAL NOTICE